Review of

Transect Sampling Methods to Obtain Population Size Estimates for Northeastern Offshore Spotted and Eastern Spinner Dolphins

Paul A Medley
Sunny View
Main Street
Alne, UK
paul.medley@virgin.net

Executive Summary

The population estimates resulting from this research are clearly a significant improvement over previous estimates. Correcting for large bias of subjective estimates and observations has improved accuracy. Rigorous statistical treatment has produced more reliable estimates. In my opinion, these population estimates are adequate for basing management decisions and setting quotas.

Simulation studies indicate there are no basic errors in coding in the multivariate transect analysis. Large fluctuations in estimates have been reduced, which is consistent with marine mammal ecology. However, the analyses undertaken are complicated, so some care will be needed to ensure final results include most recent data corrections and statistical models.

Each stage should be clearly laid out and documented. While the reviewed manuscripts covered all methods applied, not all were used in the final estimation. The actual methods applied in each year should be made clear, as well as providing confidence intervals / error estimates, as is already done. This information may be particularly important for the population model, where year estimates may require different treatment.

I have suggested a few improvements, which could be considered in future developments of the methodology. There is a tendency for overfitting to reduce apparent errors with too great a reliance on purely statistical criteria to include parameters. More careful consideration of the theoretical models could address this. I believe the greatest improvement will be obtained from further work on the use of covariates in transect sampling, and this should be the priority for future research. Otherwise, the biggest reductions in confidence intervals will probably be obtained from building conditional likelihoods between sequential year estimates, as will be done through population modelling.

Introduction

This review considers the methodology used to derive population size estimates for two species of dolphin found in the Eastern Pacific. The two species of dolphin, northeastern offshore spotted (*Stenella attenuate*) and eastern spinner (*S. longirostris orientalis*), are caught as by-catch in the Eastern Pacific tuna purse seine fishery, and are thought to be in a depleted state. The time series of population size estimates generated from this study will be used to fit a population model and will form the basis for setting advice to bring about recovery in the dolphin populations. This review only deals with the application of transects to derive population estimates, and does not cover use of the data in any population model.

The population estimate methodology divides into four activities described in four documents provided by the NMFS La Jolla Laboratory Researchers / CIE. The first two documents describe methods used to improve data accuracy. The third document describes new statistical methods to analyse the available data. The fourth document describes how the final estimates were obtained. Additional material and background information were obtained from the authors on a visit to the La Jolla Laboratory 15-17th October 2001. In particular, the authors made available data used to develop and test their methods. I found these data were particularly useful in gaining an insight into methodological issues and improving confidence in their work.

The researchers and staff at the SWFSC, La Jolla were very helpful in conducting this review.

Methodology Review

My overall judgement is that the results from the methods applied are reliable and represent a significant advance over previous methods. The manuscripts describe a considerable amount of work conducted over the last few years. The statistical approach, with some minor reservations, was statistically sound. Adequate evidence was presented to confirm there were no software coding errors or similar mistakes which would have a significant impact on the final results. The following sections review each of the documents in more detail and consider, in particular, the use of covariates in the analysis of transect data.

Kinzey, D., Gerrodette, T. Fink, D. Accuracy and Precision of Perpendicular Distance Measurements in Shipboard Line-Transect Sighting Surveys.

The analysis aimed to improve the accuracy of distance measures by applying various corrections to the angle measures taken. Two methods were applied to the horizontal angle and vertical angle (reticle) observations.

The effect of rounding horizontal and vertical angle measurements was tested by calculating the autocorrelation of frequencies which detects higher than expected frequencies at rounding values, for example values around the reticle marks. This was 'corrected' by adding random values drawn from a uniform distribution to smear values between frequency peaks. Where this technique may improve estimates with severe bias, I would have expected with normal rounding effects it will only add variance to estimates. This was confirmed by the analysis, and without good reason to alter the original observations, this technique was not used.

The second analysis tried to correct bias in distance estimates by comparing distances measured using reticles with more accurate radar distance measures. Various corrections were tried using regression and theoretical corrections for light refraction. A light refraction correction was found to be the most accurate for correcting estimates of distance to radar targets.

I support the application of theoretical correction rather than regression. Correction for well-known physical effects will improve accuracy of distance measures without the potential errors which can be introduced through statistical analysis. The most accurate correction involved pressure and temperature data which was not available for all years. This could be applied where possible, but the relative higher variance of population estimates where these data are not available should, strictly speaking, be estimated and transmitted to the final population estimates.

Gerrodette, T. Perryman, W. and Barlow, J. Calibrating Group Size Estimates of Dolphins in the Eastern Tropical Pacific Ocean.

Corrections to bias in subjective school size estimates were made based on a sample of estimates for which exact school sizes were known. Exact school sizes were obtained from aerial photographs. Estimated school sizes were calibrated by a regression of photographic counts on the school size estimates (best, high and low estimates). Separate regressions were undertaken for each observer, and factors such as sea state and year were estimated within the model, if their inclusion met the Akaike Information Criteria.

Given the sensitivity of the final population estimates to school size, and the size of bias discovered by the researchers, this technique appears to have improved the estimates greatly. This is an important improvement.

I have only a few points. The model form could be revisited. I am not sure how many different models were tried, so this suggestion may be redundant. As the aim is to obtain subjective estimates as close to the photographic counts as possible, these could be treated as the dependent Y-variable in a generalised linear model regression (GLM), rather than the independent X-variable. For example, the model in generalised linear form[1] could be:

$$E(S_{photo}) = Exp(1 + S_{best} + S_{low} + S_{high} + Beaufort * S_{best} + Year * S_{best})$$

where logarithms may, or may not, be taken of school sizes (S). Once fitted, the model allows the same treatment of all data to get expected school sizes.

In the calibration, allowance is made to obtain a weighted mean estimate from the best, high and low school size estimates on a linear scale, whereas the main regression is on a log-scale. This makes the fit non-linear. Secondly, the calibration requires reversal of the equation. This seems a little odd as the fit minimises the squared difference between the *corrected* photographic count and the subjective estimate. I would normally expect the difference between the *corrected* subjective estimate and the photographic count to be minimised through regression, as this is the final desired result.

The differences in treatment between years needs to be passed on to the analysts fitting the population model. Treatment of the underlying data, particularly calibration of the school size, was not consistent and in some years the correction was more accurate. The least accurate estimates are at the beginning of the time series, probably the least critical period for the population model. Nevertheless, the poorer calibration of the earlier part of the series should probably be reflected in fitting the population model, through statistical weighting or similar techniques.

All observers are treated separately. I suspect this has as much to do with practical problems of fitting a model simultaneously to all observers as theoretical considerations (the model would be very large). However, at least theoretically, better results could be obtained, particularly for new observers, if a single model could be fitted with parameters accounting for observer effects. Fitting separate models is equivalent to fitting a single model with full observer interaction terms and separate error scale parameters for each observer. With a single model, it may be possible to reduce this number of parameters.

Forcada, J. Multivariate Methods for Size-Dependent Detection in Conventional Line Transect Sampling.

The manuscript presented four methods to bias-correct mean group size and estimate the detection function, which is used to calculate the effective half-width of the transect. The methods tried were a simple least-squares and a robust regression for estimating the mean school size, and a multi-variate line transect analysis using parametric and non-parametric detection models.

[1] McCullagh, P. and Nelder, J.A. 1982 Generalized Linear Models. Chapman and Hall, London.

The regressions illustrated the high influence of a small number of large schools, which may introduce significant error into the population size estimates. The multi-variate line transect analysis simultaneously estimated density taking into account school size effects on probability of detection, and mean school size. The analysis found the parametric detection models gave the best trade-off between bias and precision. Finally, for estimating confidence intervals, using the bootstrap, a sample of school sizes were drawn from smoothed distribution.

The final results and conclusion seem sound. The method should account of the most important covariate effects, notably the greater detectability of larger schools. Although I believe some improvement is possible in the multi-variate model, I believe the general approach is the correct one. The analysis was carefully thought out and a variety of models were tested covering the main approaches. The methods were tested on simulated data, which demonstrates the analyses were correctly coded and should provide correct estimates if the assumptions are correct.

The main criticism of the method stems from the lack of analysis of the covariates and consideration of the theoretical influence they may have on the school detection. I believe it may be possible to make better use of the covariates and make the final estimation easier, through consideration of:

- The distance at which a school was detected. This information, I suspect, could be valuable in estimating the probability a school escapes detection at a given perpendicular distance from the transect line (see Appendix 1).
- The theoretical relationships among covariates, and between covariates and detection. Having considered possible relationships, I can see this issue would not be an easy one to solve. However, a functional start-point may be based on visual acuity measures which include distance and object size. Attributes of the school could be used to estimate its effective size (i.e. detectability), whereas visibility and swell height, might be used to estimate the effective area of search. This should at least give some sort of justification to the way covariates are used in the model, and probably lead to some improvements in estimates (see Appendix 1).

Gerrodette, T. Forcada, J. Estimates of Abundance of Dolphin Stocks Affected by the Tuna Purse-Seine Fishery in the Eastern Pacific Ocean.

Overall, these new estimates of abundance are based upon sound statistics and significant improvements in data accuracy (notably school size estimates and distance measures). The final resulting population estimates show considerable decrease in fluctuations from previous, less rigorous, analyses (Figure 1). This is to be expected, as marine mammal populations would probably not fluctuate wildly from year to year.

Of greatest concern to me was different treatment of each year's data. The same corrections could not be applied in all years through lack of data, and transect detection models in each year were chosen independently based purely on statistical criteria. While I understand the approach of treating each year as an independent observation, model structure errors might produce time series effects which might be wrongly interpreted by a population model. For example, why would there be a ship effect in some years and not others (Table 1)? Including terms without some sort of theoretical justification, (such as alterations to the observation platform), may introduce artefacts in the results. I would much prefer a single model based upon some theory which was consistent across years. However, in practice, the models chosen on the basis of purely statistical criteria show little variation.

The survey design was not random, but designed for maximum cover of the populations of interest and takes into account necessary logistical constraints. It is difficult to see how this might be improved upon. However, the non-random nature of the transect could be taken into account in the bootstrap, in particular, as sequential transect days are not independent samples (see Appendix 1).

Some issues over detectability at greater distances were removed by excluding sightings beyond a fixed perpendicular distance. Unfortunately the results appear to be sensitive to the choice of distance, but this only reflects the general problem of dealing with large influential schools. This has been handled as objectively as possible.

Figure 2 summarises the methods used to estimate the final abundance. Some techniques are described but not used. The analysis did not account for rounding errors in the vertical and horizontal angle. The regression model fitted to the expected school size was not used, but rather the weighted average based on the detection function was used instead. Documenting discarded techniques is worthwhile as it indicates what has already been considered and helps in formulating new research. However, the authors should make it as clear as possible what was used for final results, perhaps separating out discarded techniques into separate documents or appendices.

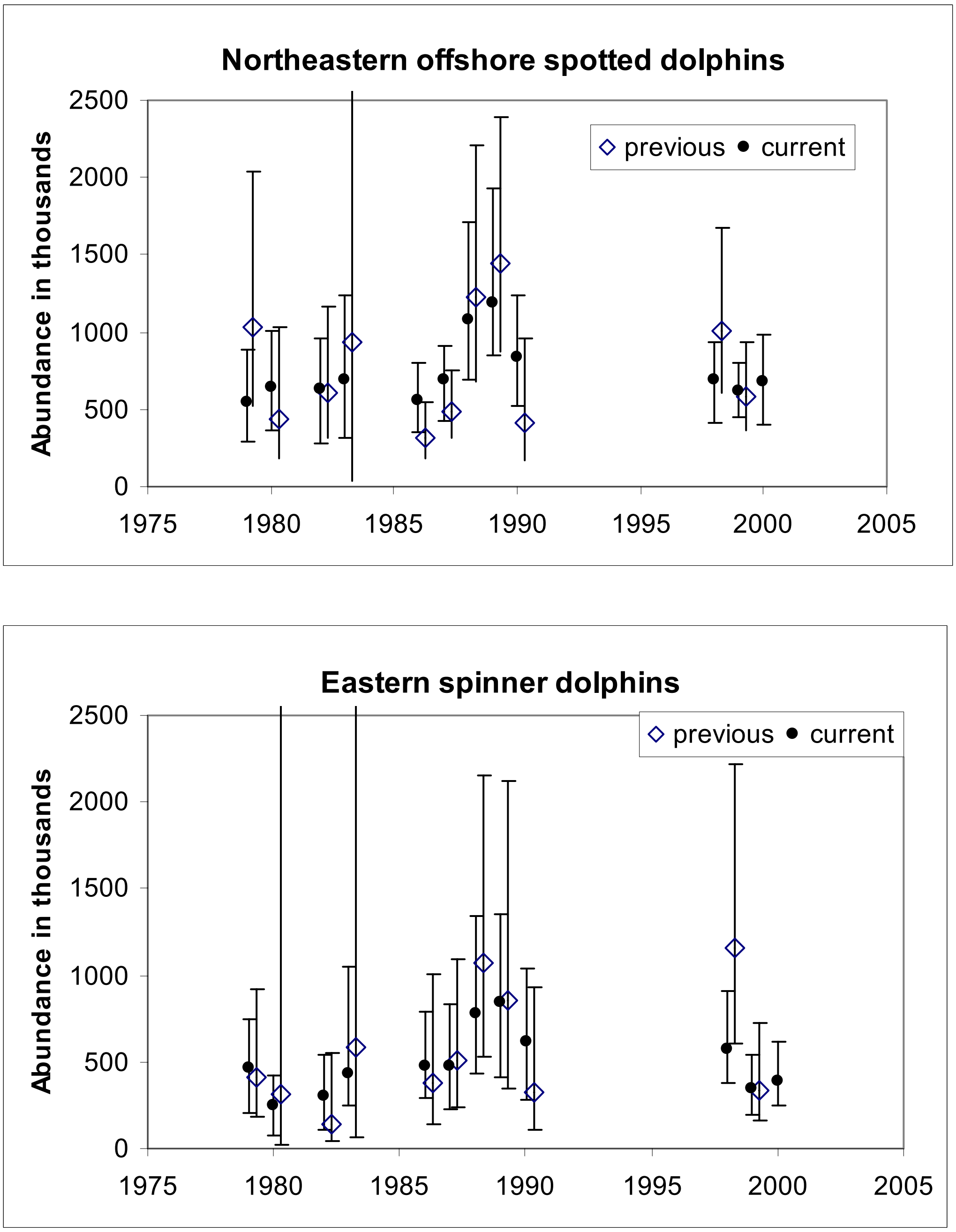

Figure 1 Previous and revised estimates of population size for the two dolphin species of interest.

Figure 2 The flow diagram summarises which techniques were applied to estimate the final population size for the northeastern offshore spotted (*Stenella attenuate*) and eastern spinner (*S. longirostris orientalis*) dolphins. The flow chart applies for each sighting category (e.g. spotted dolphins, eastern spinner dolphins, unidentified dolphins etc.). When the sighting category included unidentified dolphins, the abundance was estimated as shown, then prorated among the sub-categories. Thus, the final spotted dolphin abundance estimate is a sum of the estimate from identified spotted sightings and an estimate from the prorated abundance of unidentified dolphin sightings. Total group size was only used as a covariate to model the detection function, but the species specific group size was used to estimate individual density as well.

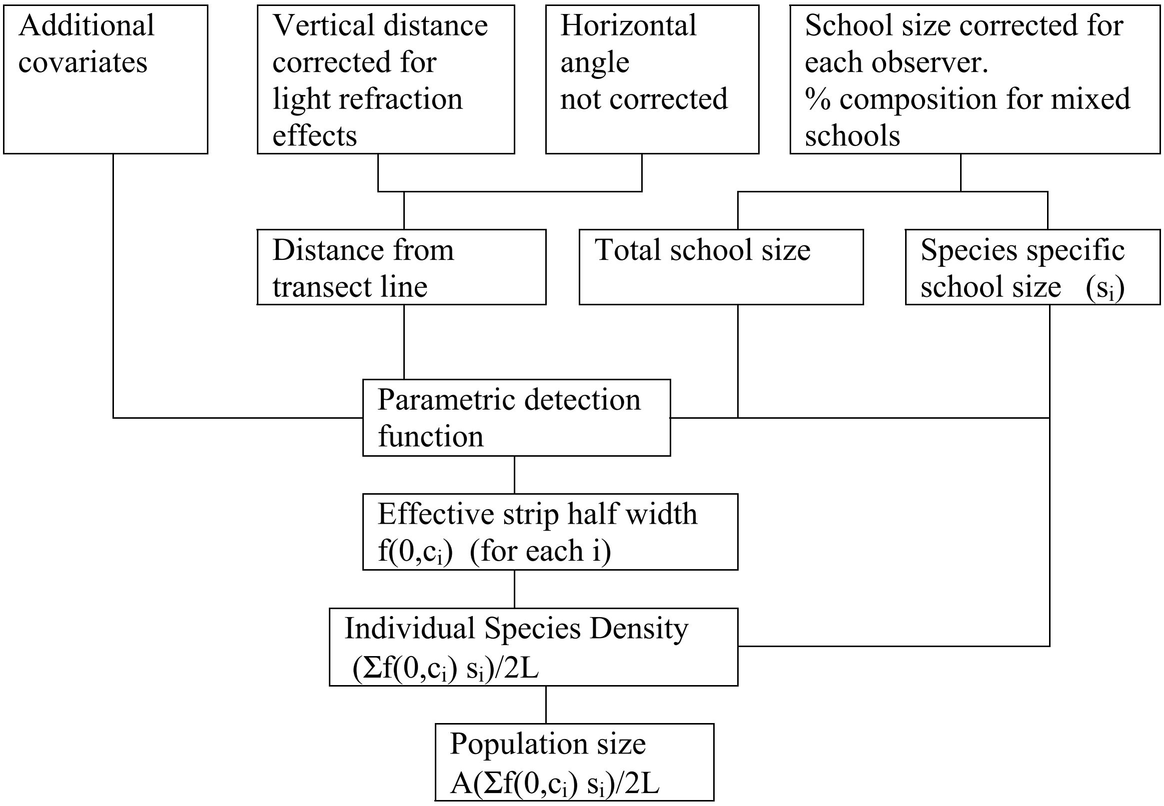

Table 1 Models and data used in population estimates by year. The scale parameter of the parametric detection function was a linear function of covariates. In practice, variation in the data and models used in each year is small.

Year	Distance Calibration	School Size Calibration	Eastern spinner dolphin		Offshore spotted dolphin	
			Model	Covariates	Model	Covariates
1979	X	X	HN	pd+gs pd	HN	pd+gs pd
1980	X	X	HN	pd+gs pd	HN	pd+gs pd
1982	T	P	HN	pd+gs pd	HN	pd+gs pd
1983	T	P	HN	pd+gs pd	HN	pd+gs pd
1986	T	P	HN	pd+gs	HN	pd+gs+ship
1987	T	T	HN	pd+gs	HN	pd+gs+bird
1988	T	T	HN	pd+gs	HN	pd+gs+ship
1989	T	T	HN	pd+gs	HN	pd+gs+ship
1990	T	T	HN	pd+gs	HN	pd+gs
1998	T	T	HN	pd+gs	HN HR	pd+gs pd+gs
1999	T	T	HN/HR HN/HR	pd+gs pd+gs+time	HN HN	pd+gs pd+gs+time
2000	T	T	HN HN	pd/pd+gs pd+bf	HN HR	pd+gs pd+tgs+sh

Calibration: X=not applied, T=Applied, P=Partial or limited application of method
Detection Models: HN=Half-normal, HR=Hazard-rate
Covariates: pd=perpendicular distance, gs=school size, tgs=total school size, bird=presence of birds, ship=ship conducting the transect, bf=beaufort scale, sh=swell height, time=time of day

Comments on Using Covariates in Transect Models

The model estimates a linear predictor from covariates recorded at the time a dolphin school is sighted. The linear predictor estimates the scale parameter for the line transect detection model, which describes the probability of detection as a function of perpendicular distance from the transect line. The aim is to adjust the detection model based on the covariates, so allowance is made for attributes of the school which may otherwise bias the results.

The main two concerns addressed by the covariate model are:

1. The fitted model may remove bias in the measure of absolute abundance caused by school attributes. Most important of these attributes is school size. If larger schools are more easily detected, results may overestimate the number of large schools there are in the population. Overestimating the number of larger schools could significantly bias the estimate of population size.
2. Trends in detectability of schools may produce false trends in population size. For example, if birds are an important visual cue, bird population size could affect detection of the schools, making estimates also dependent on seabird populations. Trends in seabirds could then be reflected in the results.

The inclusion of covariates are an important improvement in the model. However, the method adopted raises two related issues which I believe are significant. Firstly, the functional form of the detection model. Whereas without covariates, choice of the detection model is probably less critical, I would believe effective use of covariates requires accurate modelling of scale and form of the relationship. There appears to be no standard approach in the analysis of line transect data.

The second issue is the change in model for each year. It may well be that covariates have different effects in different years, and therefore the model may effectively change. However, there is a potential problem with different treatments of the data in each year, as it is possible that model structural errors may exaggerate changes in population sizes between years. For the current analysis, my judgement is that this issue is not of great concern as:

1. The functional form of the fitted model changed very little over the years.
2. The full fitting process was included in the bootstrap, so potential variation from models is represented in the confidence of the population estimates.
3. I presume these estimates will be used to fit a population model. The effect of introduced annually independent errors will be reduced through this fitting, as long as they are not very large.

Nevertheless, I would recommend in future more consistent treatment between years. Although I understand that fitting models between years may be perceived as producing more accurate estimates, this may not be the case. With smaller data sets in each year, there is an increased chance of type I and II errors, and structural errors. Hence, I would tend to apply the same model across years unless there is evidence that school detectability has changed. I cannot see why the detection function would be different between years, although I can see there is no good reason to choose one function over another apart from likelihood. I can also see why the affect of school attributes might be different from year to year. Weather conditions, bird and dolphin behaviour may change annually. Nevertheless, I would place more onus on testing the differences between years as opposed to assuming there are genuine differences by default.

I would recommend choosing a consistent function which may be applied to all years. This would allow parameter estimates and observation errors to vary among years, but there would be no differences in structural errors. As I cannot think why the model should vary annually, I would presume this would produce more accurate results. In addition, given a single model the application of the analysis would be simpler.

Finally, I would suggest that fitting a covariate, such as school size, should be applied even for those years it is not statistically significant. It ensures the effect of the covariate is adjusted for, as far as possible. 'Overfitting' in this multi-year context is not a critical problem as individual parameter estimates are not used, only the final predictor which should be less sensitive to overfitting problems. For example, correlation values close to 1.0 in the parameter correlation matrix may produce poor estimates for the individual parameters, but these should cancel out in producing the estimated scale parameter for the detection function.

A further refinement, suggested by Jaume Forcada, was obtaining a population estimate from the average estimated population sizes weighted by their likelihood. This would further reduce differences between years. I would still prefer a single, justified treatment for all years as it would be simpler and less prone to structural error.

Future research might consider addressing how covariates could be used more effectively by collecting experimental data specific to this task. An obvious variable under control of the researcher is the speed of vessel. Varying the speed should vary the detection function, and hence provide information on its form. Other techniques, such as using the helicopter for spotting, and so on, might also address this problem.

Appendix 1 Some Exploratory Analyses of the Sightings Data 1998-2000

These analyses aimed to check some results from the full analysis under review and indicate some possible ways to improve it in future. The data were kindly provided by Tim Gerrodette and Jaume Forcada for school sightings 1998-2000. There were some inconsistencies in the data sets provided, probably due to variations in corrections applied. Furthermore, I derived some data, such as days effort, which are probably not exact. These data should be adequate for this superficial analysis, but should not be used as final results.

Table A.1 gives the distribution of schools within each search day. The distribution is overdispersed relative to the Poisson, but presents no special problems for the techniques applied.

Table A.1 Distribution of the number of schools sighted within each day of search effort. The observed distribution is over-dispersed with respect to the Poisson distribution. In addition, there is presumably a truncation effect, as spotting a school requires the vessel to go off-effort as it estimates its size.

Bin	*Frequency*	Poisson
0	592	454.72
1	157	303.32
2	63	101.16
3	33	22.49
4	18	3.75
5	11	0.50
6	7	0.06
7	3	0.01
8	1	0.00
9	0	0.00
10	0	0.00
11	1	0.00

Table A.2 School sightings on sequential days using data 1998-2000. The approximate G-test (Sokal and Rohlf 1981, pp 737-738) confirms sightings on sequential days are not independent (G_{adj} = 96.76, is much greater than $\chi^2_{.001[1]}$). These results are not species specific.

		Next Day		
		No Sightings	**Sighting**	**Total**
Previous Day	**No Sightings**	459	130	589
	Sightings	129	162	291
	Totals	588	292	880

Whether schools are sighted depends upon whether at least one school was sighted the previous day (Table A.2). This occurs because schools aggregate and transects are not random. This is a fairly strong effect, and essentially reduces the apparent sample size (i.e. variance), although I would not expect the effect to be very large because the sample size is large. Its effect could be explored through simulation, although I would not expect the best estimate to change much. The autocorrelation could be included in an empirical bootstrap by sampling using a Markov chain from effort-day observations grouped into those which had sightings on the previous day and those which did not. More complex conditional sampling could be developed if the issue was found to contribute significantly to the estimate variance.

Modelling with Covariates

A general form of a detection model describes a Poisson process where the probability a school is seen depends upon the time it is within the observer's field of vision

$$\Pr(\text{School detected}) = 1 - e^{-\int P(\beta,t)dt}$$

If the movement of the school is small relative to the movement of the vessel (as supported by mapping the position of the schools relative to the movement of the vessel), the school may be detected at any moment over a distance *f*, which starts when the school enters the maximum visible range and passes the 90° limit angle of observation (Figure A.1).

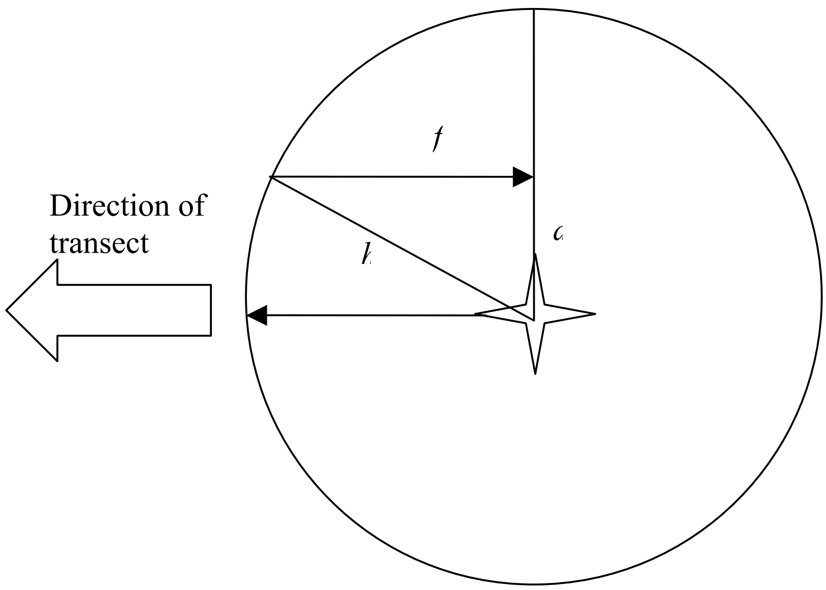

Figure A.1 Diagram illustrating distance (and time taking into account vessel speed) over which a school might be detected, where school movement in relation to the transect line is effectively zero. The distance of detection, *f*, depends on visibility or horizon (*h*) and perpendicular distance from the transect (*d*).

The function, $P(\beta,t)$, is the probability a school is detected as a function of distance, time and attributes of the school (such as school size). For example, whether the school is noticed may be proportional to the angle made by the object with the observer's visual field.

$$P(\beta,t) \propto \text{Arctan}\left(\frac{a}{\sqrt{d^2 + (f - vt)^2}}\right)$$

where *a* is the half-width of the school. It is very likely that whether an observer notices an object will be a non-linear function of this visual field angle, but may not decline as rapidly as the half-normal (Figure A.2).

Even if a complex function is identified as the most appropriate, a simple approximation would probably be adequate. However, I suspect any demonstrable improvement of the hazard rate, half-normal or other tested models could only be derived from experimental data. For example, varying vessel speed, distributing test targets for trial transects, or using the helicopter to search and monitor schools (as has already been done). This could be used to support a particular detection function.

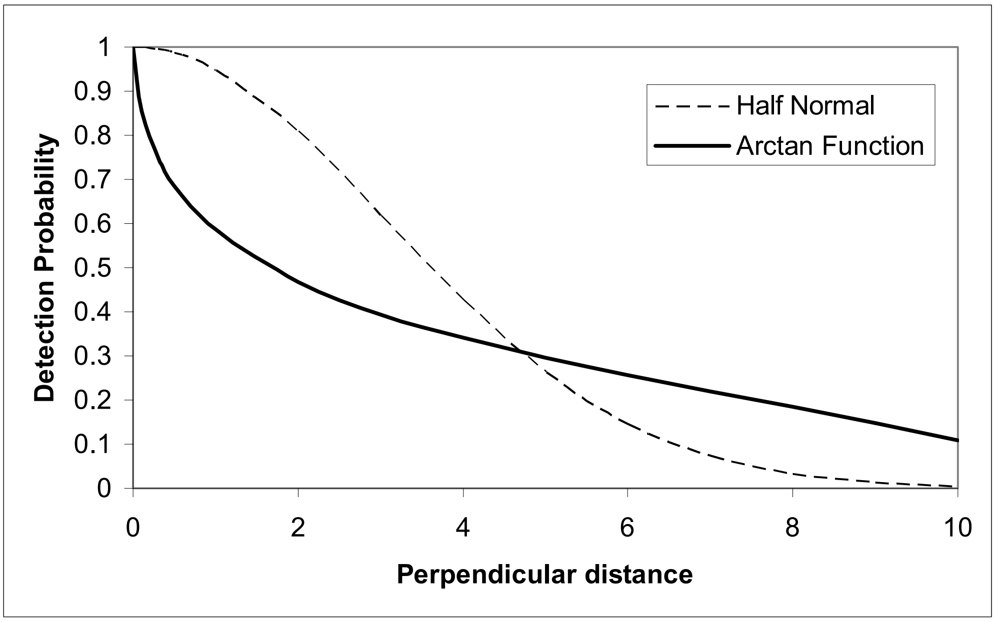

Figure A.2 The visual field angle shows a sharper decline in detection than the half-normal, for example. By itself, the Arctan function does not appear to be realistic, and the models used (hazard rate or half-normal) as good as any.

Relationships between covariates could probably be considered on a theoretical level. I am unclear on exactly what relationships they could be. Visibility reflects the maximum distance a school might be observed. The size of school would affect the number of splash events which might be noticed. The swell height might make fewer such cues visible, decreasing the effective school size. The sun position could affect the search area, effectively removing a segment of the ocean available to search when the vessel is heading into the sun. Birds make schools more noticeable when associated with schools, but may help hide schools when not associated as they attract observers' visual search effort. Windspeed may produce white-tops which might make cues more difficult to spot. Different species in mixed schools may be more visible than others.

It may well be worthwhile developing a model with observers who may be able to indicate what difference these or other cues make. In modelling terms, they can probably be divided into:

- Those that change the effective size of the object (e.g. school size, swell height)
- Those that change the effective search area (e.g. visibility, sun position)

A parallel statistical approach looks less promising. Conceptually, covariates are there to measure some non-observed variable, "detectability", which might be derived from a factor analysis. A Principle Components Analysis did not yield useful results, although there were clear correlations between covariates. Most notable was common weather effects, from which might be derived a poor to good weather variable (Table A.3). However, for this approach to be useful, observations would have to be recast so that their relationship to detectability would be clearer.

Table A.3 PCA components for standardised weather variables. The high variance which can be allocated to the first component suggests a strong relationship between variables, which may be more useful in modelling detectability, or least making models more parsimonious.

Components	1	2	3	4	5
Beaufort	-0.64	-0.37	-0.35	-0.27	-0.51
Swell Height	-0.62	0.16	0.76	0.07	0.10
Swell Direction	0.30	-0.64	0.43	-0.55	0.08
Wind Speed	-0.30	-0.44	-0.27	0.24	0.76
Wind Direction	-0.15	0.48	-0.21	-0.75	0.38
Loadings	92%	4%	2%	2%	1%

Finally, another approach which I believe may prove useful is to model the actual distance when a school is observed, given its perpendicular distance, the set of covariates and that it is observed. The model I set out here is fairly crude, but a refined version may better use of data in formulating a detectability function.
The start point is to account for possible distances based upon the perpendicular distance. The average distance of detection is:

$$c = E\left(\text{Detection distance} \mid d\right) = \frac{\int_0^{f/v} \sqrt{d^2 + (f - vt)^2}\,dt}{\int_0^{f/v} dt}$$

$$= \frac{d^2}{2f}\left(\frac{fh}{d^2} + \ln\left(\frac{f+h}{d}\right)\right)$$

where

$$f = \sqrt{h^2 - d^2}$$

h=maximum distance (visibility) and *d*=perpendicular distance from the transect line, v=velocity of the vessel.

The use of average distance assumes detection is equal for all distances for each given perpendicular distance. This is clearly unlikely, and observed distances are likely to be biased to shorter values.

Otherwise covariates considered were total school size, visibility, swell direction and height, vertical and horizontal sun angles, wind speed and direction, Beaufort scale, whether birds are present and whether the school contained more than one species. The actual cue used to spot a school was not used, as it is an attribute of the sighting event rather than the school or searching conditions.

The model was a simple multiplicative model:

$$\mu_i = Exp\left(a_0 + \sum_j a_j x_{ji}\right)$$

$$\text{minimise} \quad \sum_i \left(\mathrm{Ln}(y_i) - \mathrm{Ln}(\mu_i)\right)^2$$

where y_i is the observed distance at which the school i was first sighted, a_j refers to the fitted parameter to minimise the log-normal log-likelihood for covariate j, and x_{ji} the covariate j for sighting i.

The majority of covariates did not explain much variation in the observation distance (Table A.4). The most important variables, notably school size and (perpendicular) distance were used in the parametric detection model, and so these results concur with those for the reviewed analysis. However, visibility and the presence of birds also turned out to be potentially important variables, as did the interaction term between birds and school size. Birds are an important factor in identifying schools, but may be more important for small rather than larger schools. Full interaction terms were not explored, but it seems likely that there will be other interactions between visual cues and the size of the school.

Table A.4 Analysis of variance table for the multiplicative regression fitted to distance of sighting. Remaining covariates explained only small proportions of the variance, although models investigated were not exhaustive.

	Sum of Squares	Degrees of Freedom	Mean Square	Approximate F Ratio
Constant	270.21	1		
Ln(Distance)	237.65	1	32.56	106.03
Ln(School)	204.49	1	33.16	107.99
Ln(Visibility)	188.18	1	16.31	53.11
Birds	179.31	1	8.87	28.89
Birds*Ln(School)	173.53	1	5.78	18.84
Error		584	0.31	

Although the model appeared reasonable, it was limited in the proportion of variation it could explain (Figure A.3). In particular, some observed values were unexpectedly low, and these influence the results. The reason why some schools are not spotted until close to the vessel cannot be explained with the current set of covariates.

By fitting an appropriate model to the sighting distance, it should be possible to obtain the stochastic process model, $P(\beta,t)$, which can be then used obtain the probability of detection based on the perpendicular distance from the transect line through integration. This approach may be more accurate than trying to fit the detection model to the perpendicular distance directly.

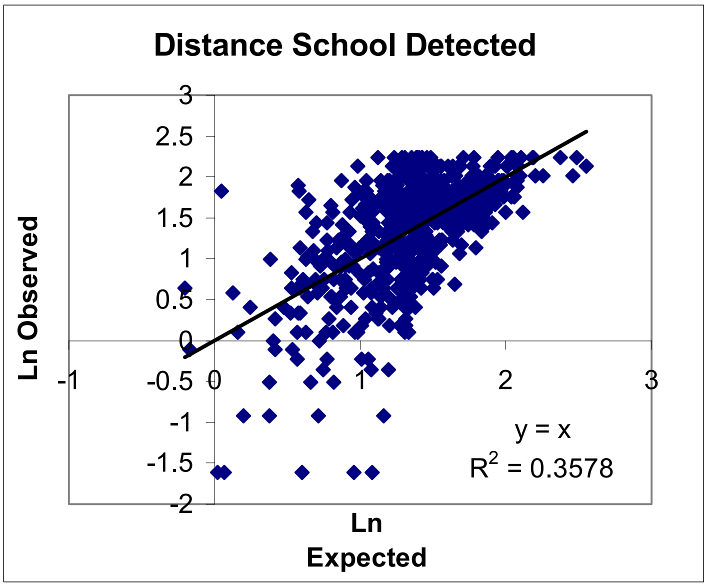

Figure A.3 Observed and expected distances to schools when they are seen. The log regression model has a low R^2, although a detection model based on perpendicular may remove some of the remaining unexplained variation.

Appendix 2 STATEMENT OF WORK

Consulting Agreement Between The University of Miami and Dr. Paul Medley

December 17, 2001

General

The topic of the review is the abundance of several species of tropical pelagic dolphins that associate with tuna and are killed in the eastern tropical Pacific purse-seine tuna fishery. In 1997, the US Congress proposed changing the definition of "Dolphin-Safe" tuna, but it made the change in definition contingent on the results of studies of the impact of the tuna fishery on depleted dolphin populations. Estimates of dolphin abundance based on cruises carried out in 1998-2000 form a central part of these studies. The tuna-dolphin issue is a controversial issue among NMFS, US tuna industry, foreign tuna industry, and environmental groups.

The main task of the consultant is to review the methods used to estimate abundance from line-transect data, including covariate detection models. The fact that these dolphins occur in a wide range of school sizes presents unique problems for the estimation of expected group size, so considerable effort has been devoted to this analysis. The expertise of the consultant should include knowledge of statistics and methods of population estimation, especially distance sampling (line-transect) methods.

Documents supplied to the reviewers will include draft manuscripts describing the covariate analysis, simulations to test the performance of several estimators, calibration of school size estimates, and assignment of partially identified sightings. Background papers will include previous relevant publications and reports. The raw data and software used in the analysis will be available to the reviewers if they wish.

Specific

The consultant's duties shall not exceed a maximum total of 2 weeks- several days to read all background documents, attend a three-day meeting with scientists at the NMFS La Jolla Laboratory, in San Diego, California, and several days to produce a written report of the findings. It is expected that the individual contribution of the consultant shall reflect the consultant's area of expertise; therefore, no consensus opinion (or report) will be accepted. Specific tasks and timings are itemized below:

1. Read and become familiar with the relevant documents provided in advance to the consultant;

2. Discuss background documents with scientists at the NMFS La Jolla Laboratory, in San Diego, CA, for 3 days, from October 15-17, 2001;

3. No later than November 16, 2001, submit a written report of findings, analysis, and conclusions. The report should be addressed to the "UM Independent System for Peer Reviews, " and sent to David Die, UM/RSMAS, 4600 Rickenbacker Causeway, Miami, FL 33149 (or via email to ddie@rsmas.miami.edu).

Signed________________________

Date____________

Appendix 3 Bibliography and Materials Provided for Review

Kinzey, D., Gerrodette, T. Fink, D. Accuracy and Precision of Perpendicular Distance Measurements in Shipboard Line-Transect

Gerrodette, T. Perryman, W. and Barlow, J. Calibrating Group Size Estimates of Dolphins in the Eastern Tropical Pacific Ocean.

Forcada, J. Multivariate Methods for Size-Dependent Detection in Conventional Line Transect Sampling.

Gerrodette, T. Forcada, J. Estimates of Abundance of Dolphin Stocks Affected by the Tuna Purse-Seine Fishery in the Eastern Pacific Ocean.

The authors also provided sundry materials as background, which are presented in this report, as well as the sightings data.